

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>6</sup> : G06F 13/00, 15/00, 17/22, 17/30, G06T 11/00		A1	(11) International Publication Number: <b>WO 99/49396</b>
			(43) International Publication Date: 30 September 1999 (30.09.99)
(21) International Application Number: PCT/US99/06333		(81) Designated States: AU, CA, IL, JP, MX, NO, NZ, RU, SG, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).	
(22) International Filing Date: 23 March 1999 (23.03.99)		Published <i>With international search report.</i>	
(30) Priority Data: 09/046,364 23 March 1998 (23.03.98) US			
(71) Applicant: WEBIVORE RESEARCH, LLC. [US/US]; One Brook Street, Watertown, MA 02172 (US).			
(72) Inventors: LENK, Cecilia; 81 Boyd Street, Watertown, MA 02172 (US). GOULDSTONE, James, B.; 506 Washington Street, Brighton, MA 02135 (US). REIDY, James, G.; 9 Lincoln Street, Westboro, MA 01581 (US). GORT, Carl, J.; 39 Royal Crest Drive #7, Marlboro, MA 01752 (US).			
(74) Agent: HENNESSEY, Gilbert, H.; Fish & Richardson P.C., 225 Franklin Street, Boston, MA 02110-2804 (US).			
(54) Title: INTERNET RESEARCH TOOL			
(57) Abstract			
<p>A system for conducting research on the Internet accesses a pre-selected web site on a computer network through a host site and provides a view of a page from the web site on a computer monitor in an environment downloaded at least in part from the host within which data can be clipped from the view and saved to a file, along with site-specific reference data, which includes, for example the site name, author, and URL. A human-perceptible document can be produced from the clipped data and the data can be edited.</p>			

***FOR THE PURPOSES OF INFORMATION ONLY***

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	BS	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Larvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TC	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Gulaca	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NB	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

WO 99/49396

PCT/US99/06333

- 1 -

INTERNET RESEARCH TOOL

## REFERENCE TO MICROFICHE APPENDIX

- A microfiche appendix is attached to this application. The  
5 appendix, which includes a source code listing of an embodiment relating to the invention, includes 166 pages of source code on 3 microfiche.
- A portion of the disclosure of this patent document contains material that is subject to copyright protection. The copyright owner has no objection to the facsimile reproduction by anyone of the patent document or the patent disclosure as it  
10 appears in the Patent and Trademark Office file or records, but otherwise reserves all copyright rights whatsoever.

## BACKGROUND OF THE INVENTION

- This invention relates to an Internet research tool.
- The Internet enables users to access an abundance of information from  
15 computer systems throughout the world. Users can find useful information pertinent to their interests by using a browser and a computer-implemented search program known as a search engine. A search engine reads keywords that a user enters into a user interface, locates the keywords in documents published on a computer running a World Wide Web server process or a computer supporting Usenet, and provides  
20 access to the documents containing the keywords.

- With regards to the World Wide Web, when an interesting World Wide Web site is located, an individual can save data that refers to the location of the site so that he or she may return to the site at a later time. Netscape Navigator<sup>TM</sup>, a product from Netscape Communications Corporation, provides "bookmarks" for this purpose.  
25 Bookmarks are stored in a personal file and are essentially addresses to sites.

The amount of information retrieved by a search engine can be overwhelming and of minimal quality to the person who initiated the search. Due to

WO 99/49396

PCT/US99/06333

- 2 -

the enormous amount of information retrieved and the desire for selectivity, software is available that locates information on the Internet that may have value to a group of individuals.

5

#### SUMMARY OF THE INVENTION

The present invention features a computer-implemented method for conducting research on the Internet. The method accesses a web page on a computer network and provides a view of the web page on a computer monitor. A portion of 10 the web page can be clipped from the view and automatically saved to a file. The method annotates the file and produces a human-perceptible document from the clipped data. The human-perceptible document is output, for example, to a computer monitor.

A GUI element switches the research mode from a search mode to a collect 15 mode, including distinct sets of tools used in the respective modes. In the collect mode, the method enables a user to gather information from a selected part of a web page, including text, images, audio data, and other multi-media data. The gathered information is saved to a clipping file on disk.

Annotation is included in the same clipping file as the collected 20 information. The annotation is system-provided and automatically saved with the clipping. This system-provided annotation may be bibliographic information (e.g., in a cite format) and may include a document name, an author, a URL, and a date. Additional annotations, composed by a user, can be included in the same file.

Sites are pre-selected to control the quality of the content retrieved. The 25 method presents the pre-selected sites beneath a hierarchy of topics. After a pre-selected site name is selected and prior to accessing the corresponding web site, the method automatically displays a preview of the site so that a user can decide whether to visit the site. A hierarchical path of topics used to reach the pre-selected sites is continuously displayed in the user interface during the search mode.

30 The method provides an efficient means to conduct research on the

WO 99/49396

PCT/US99/06333

- 3 -

Internet for professional or educational purposes, by organizing and presenting topics, locating data on a computer network that is relevant to a selected topic, and collecting selected data found on a remote computer to a local computer medium.

5 An Internet research tool 102 is capable of collecting only a portion of a web page rather than having to copy an entire web page, recognizing that not all information located by the researcher is valuable or necessary to the research. In the context of the Internet, which makes available enormous amounts of information, an Internet research tool that helps a person focus on only important items of information  
10 and gather only pertinent information, thus improving the efficiency of that person as a researcher and providing economic benefits as well, for example, reducing the amount of time spent using resources necessary to use the Internet.

#### BRIEF DESCRIPTION OF THE DRAWINGS

The foregoing features and other aspects of the invention will become  
15 more apparent from the drawings taken together with the accompanying description,  
in which:

FIG. 1 is a diagram of a computer platform suitable for supporting an Internet research tool in accordance with the invention;

FIG. 2 is a diagram of a hierarchy of topics;

20 FIG. 3 is a diagram showing a switch between search and collection modes;

FIG. 4A is an illustration of a prompt, pull-down menus, and a scrolling list;

FIG. 4B is an illustration of a site preview in search mode;

25 FIG. 4C is an illustration of a web page and search tools;

FIG. 4D is an illustration of a web page and collection tools;

FIG. 4E is an illustration of a web page and collection tools;

FIG. 4F is an illustration of a human-perceptible clipping; and

FIG. 5 is a diagram of a pre-selected web site with respect to search and

WO 99/49396

PCT/US99/06333

- 4 -

collect modes.

#### DETAILED DESCRIPTION

FIG. 1 shows a computer platform 100 suitable for supporting one or more

- 5 Internet research tools 102. The computer platform 100 includes a digital computer 104, a display monitor 106, a keyboard 108, a mouse or other pointing device 110, and a mass storage device 112 (e.g., hard disk drive, magneto-optical disk drive, or floppy disk drive). The computer 104 includes memory 120, a processor 122, and other customary components, such as, memory bus and peripheral bus (not shown).
- 10 The computer 104 has a network interface 124 to communicate with remote computer systems, such as a web server 126.

An Internet research tool 102 enables a user to conduct research on the Internet, using World Wide Web pages as the source material for the research.

- (World Wide Web sites and World Wide Web pages will be referred to simply as web sites and web pages, respectively.) Internet research tool 102 is accessed via a web browser such as Microsoft Corporation's Internet Explorer™ or Netscape Communication Corporation's Navigator™. After starting a browser that displays a browser window on a local computer, a user begins using an Internet research tool 102 by going to the web site that provides the Internet research tools 102, for example, by entering the domain name www.webivore.com beside the "Location:" prompt in the Netscape Navigator™ browser window. Thus, window 132 is a web browser window that initially displays a page from the web site that provides Internet research tools 102. Gaining access to an Internet research tool 102 may require keying in a valid user name and password before the research tool 102 displays a web page on the local platform 100.

The first web page displayed by the Internet research tool presents sets of topics in a hierarchical fashion to the user. The user selects a topic and can access Internet sites containing information on the selected topic via a hypertext link to another web site. After accessing an Internet site, the user can gather information

WO 99/49396

PCT/US99/06333

- 5 -

from the selected site by selectively saving useful information.

- Referring to FIG. 2, each Internet research tool 102 is designed for a particular set of topics. Each set of topics has a hierarchical arrangement. For 5 example, one Internet research tool contains topics for college students such as arts, biology, earth and space science, history, governments, and literature. These topics are further divided into sub-topics. For example, the Earth and Space Science topic 4 has sub-topics that include Astronomy and Space 6, Environment, and Geology. 10 Astronomy and Space 6 is further divided into sub-topics, such as a topic titled Our Solar System 8. Other Internet research tools 102 provide topics for high school students, elementary school students, and people with specific interests such as 15 interests in medical research.

Each Internet research tool 102 provides access via hypertext links to pre-selected web sites. Upon linking to a pre-selected web site, research tool 102 retrieves 15 system-provided information from the web server 126. System-provided information may be site references that include authorship and a uniform resource locator (URL) that references data on the Internet. The system-provided information is stored in a data buffer named collectflash. In the collectflash buffer, the system-provided 20 information is automatically combined with information collected from a web page at the pre-selected web site, such as text, images, audio data, or any mime-type encoded into the system. Additionally, the user can produce information, such as descriptive notes, to store with the collected data. The information from the web page at the pre-selected web site, the system-provided information, and the user-provided information 25 are saved to a storage device, which will be discussed. The file in which the information is saved is in a user-named file folder located on a local storage device.

- Referring now to FIG. 3, an Internet research tool 102 divides the research task into two primary modes, search mode 12 and collect mode 14. Each mode is displayed as a web page. Search mode 12 has distinct components from collect mode 14. Search mode 12 enables a user to search for relevant Internet sites, whereas 30 collect mode 14 enables a user to collect data from an Internet site. The position of a

WO 99/49396

PCT/US99/06333

- 6 -

graphical user interface (GUI) element switch 20 determines whether the search mode 12 or the collect mode 14 is active. The user moves GUI element switch 20 from the search position to the collect position using the mouse 110. In search mode 12,

- 5 Internet research tool 102 can access a remote Internet site, such as Internet site 16. Both search and collect modes have a unique set of tools that are displayed in a frame 26b (shown in FIG. 4A) in window 132 on display monitor 106. In collect mode 14, Internet research tool 102 can collect data from remote Internet site 16 and produce clipping file 18 on computer storage medium 112.

10 Clipping file 18 contains a user-selected portion of a web page. System-provided information 19, such as a citation, is automatically included in the clipping file 18. Additionally, user-generated information 17, such as notes, can be included with the clipping. An Internet research tool 102 is self-contained in the sense that it does not require a word-processing application to run concurrently in another window  
15 to provide functions, such as operations that paste data from a clipboard and save a file to disk. One context is set up for a view on local display monitor 106. Searching for information on the Internet, collecting information from an Internet web page, generating a clipping with user-provided and system-provided information, and saving the clipping to a local storage device takes place within the same user-interface  
20 environment.

Shown in FIG. 4A, in the frame 26b in the right portion of the window 132, are the search tools. Topic selection search tools include pull-down menu 22a, and pull-down menus 22b, 22c (shown in FIG. 4B) that enable a user to conduct a hierarchical search for a web site. The hierarchical search has three levels, each level  
25 further narrows the search topic. Using the pull-down menus, a user can select a new topic at any level at anytime while Internet Research tool 102 is in search mode 12, so long as a higher level topic is chosen prior to selecting a lower level topic. While searching for and examining web sites, the user is always aware of the hierarchical path taken to reach the web site because the pull-down menus 22a, 22b, 22c continue  
30 to display the names of the selected topics.

WO 99/49396

PCT/US99/06333

- 7 -

A prompt 21 directs the user to choose a first-level topic using the top-level pull-down menu 22a. As is typical with pull-down menus, menu selections appear on the menu when the menu is pulled down. Here, first level topics are listed 5 when pull-down menu 22a is pulled down. After choosing a first-level topic, such as the Earth and Space Science topic, the name of the topic is displayed on the top-level pull-down menu 22a and the prompt 21 directs the user to select an area within the first topic by using the second-level pull-down menu 22b (shown in FIG. 2B).

- Second-level topics beneath Earth and Space Science include Astronomy 10 and Space, Environment, and Geology. The names of the second-level topics are listed on the pull-down menu 22b when the user pulls down menu 22b. After selecting a second-level topic in the topic hierarchy, the name of the topic is displayed on the pull-down menu 22b and the prompt 21 directs the user to further refine the search topic. The topic titled Our Solar System belongs to the third hierarchical level 15 and, as shown in FIG. 4A, was selected as the third-level topic because the title appears on the third-level pull-down menu 22c. Note that the topic hierarchy does not need to be symmetrical. A first-level topic may have only one subordinate level. The topic hierarchy is continuously displayed while Internet research tool 102 remains in search mode.
- 20 Available sites are displayed below the third pull-down menu as a result from the topic selection in the third pull-down menu. The sites are presented to the user as a scrolling list 25 in the user interface. Only pre-selected sites that have gone through a screening process are visible using the scrolling list 25 and initially accessible via search mode 12.
- 25 Pre-selected sites provide for a level of quality control in the research process. For example, prior to an Internet research tool 102 allowing access to a web site by choosing the site in the scrolling list 25, the site can be checked to ensure that it is reachable and up-to-date. Additionally, the quality of the site can be measured against agreed upon expectations. For example, a site's information content can be 30 checked to ensure that it meets a determined level of scientific or historical scrutiny.

WO 99/49396

PCT/US99/06333

- 8 -

Referring now to FIG. 4B, each pre-selected web site has an associated abstract, called a preview, that briefly describes the web site. When a web site is selected from the scrolling list 25, the preview is displayed automatically in target frame 26a, which initially occupies most of window 132 in the left portion of window 132. The preview describes the content of a pre-selected site, and contains a title 27 that may be the site name, author name 28, the date the site was reviewed (not shown), and one or more category classifications (not shown). The preview enables a user to investigate a web site before accessing the site and collecting data. The previews are stored in a database on web server 126 and accessible when GUI element switch 20 is set to the SEARCH position.

Additionally, icons in a preview indicate a site's features, which may include images, audio data, video data, and other multi-media types of data. Icon 30a in FIG. 4B indicates to the user that the site has images. An icon 30b displaying a moving image inside a graphic resembling a video monitor indicates video features, and an icon 30c resembling headphones indicates audio data. The icons 30a, 30b, 30c are useful not only as visual indicators of a site's features, but also to alert the user that a time lag may occur when loading data and to alert the user that memory constraints may be noticeable when loading and saving some of a site's features.

The pre-selected web site is accessed through a hypertext link when the user selects the underlined portion of title 27 that is the hypertext link from target frame 26a. When the user clicks on the underlined portion of title 27 using a pointing device such as a mouse 110, the Internet research tool 102 replaces the preview in target frame 26a with contents provided by the pre-selected web site linked through title 27. Different pages of a web site are perused in the same manner as a user peruses web pages using the installed browser. Any web sites linked to the displayed site (e.g., through a hypertext link) are also viewable. The linked sites are not pre-selected sites so the linked sites do not have the benefit of a preview. However, the research tool 102 allows for the collection of information from any site that is displayed in the target frame 26a.

WO 99/49396

PCT/US99/06333

- 9 -

Referring now to FIG. 4C, a web page from a pre-selected web site is shown. Research tool 102 provides a consistent environment when searching for a web site and collecting data from a web site. The web page displayed in target frame 26a in the left portion of the user interface window 132 is initially displayed while in search mode and remains after the transition from search mode to collect mode.

Referring now to FIG. 4D, setting GUI element switch 20 to the COLLECT position enables research tool 102 to collect data from the web page and produce a clipping file 18. After setting GUI element switch 20 to COLLECT, the 10 web page remains in target frame 26a; however, a distinct set of collection tools are displayed in frame 26b, replacing the search tools.

Collect mode 14 provides a collect area 31 to which selected items, such as text and images, are dragged and dropped using mouse 110. Each collected item can be stored with user-generated information 17 and system-provided information 19 as a 15 separate clipping file 18. Text is gathered using copy and paste operations.

Generally, the implementation of drag and drop operations and copy and paste operations depend on the installed browser's capabilities. In collect mode 14, Internet research tool 102 produces an HTML file in which data is stored on disk. The HTML file is stored in a file folder having a filename that a user specified in a dialog box 20 when data is saved to a particular clipping.

Referring now to FIG. 4E, Internet research tool 102 displays information about clipping file 18. Clipping file 18 may be text, a graphics image, video data, audio data, or any mime-type that the system is configured to recognize. Icon 34 indicates the clipping type. For example, one graphic represents an image, another 25 graphic represents text. References, such as site name 35, site author 36, URL (not shown), and dates are automatically stored with clipping file 18. These references provide the basis for bibliographic and citation information. Additionally, the user can add notes 37 to clipping file 18.

Referring now to FIG. 4F, a human-perceptible form of the clipping file 30 18' is shown. The clipping in a human-perceptible form 18' is analogous to a page in

WO 99/49396

PCT/US99/06333

- 10 -

a researcher's notebook. Internet research tool 102 displays the human-perceptible form of the clipping file 18' in the same target frame 26a as the corresponding preview and the corresponding web site. The human-perceptible form of the clipping file 18'  
5 is displayed after the user selects the VIEW CLIP button (not shown). User added notes and a filename may be edited by choosing the EDIT CLIP button 39. Additionally, the human-perceptible form of the clipping file 18' can be output to a hard-copy device using the browser's print command.

Referring to FIG. 5, a diagram of a pre-selected web site with respect to  
10 search and collect modes is shown. Search mode 12 and collect mode 14 are implemented as HTML (hypertext markup language) files named search.html and collect.html, respectively. GUI element switch 20 is implemented as a Java applet named Switcher.java 20', which is a program written in the Java programming language and executed by a browser, such as Internet Explorer™ version 3.0 or  
15 higher, or Netscape Navigator™ version 3.0 or higher. Switcher.java 20' sets the Internet Research Tool 102 to search mode or collect mode when Switcher.java 20' detects a mouse click event.

When search.html is active, a user can choose a topic and sub-topics (step 42). As discussed, Internet research tool 102 has a hierarchy of topics. Research tool  
20 102 has a category navigator, implemented as a Java applet, to access the topic hierarchy via pull-down menus 22a, 22b, 22c shown in FIGS. 4A, 4B, and 4C. Once a top-level topic is selected, sub-topics can be selected and displayed via second and third-level pull-down menus 22b, 22c. To list the sites, the user uses the mouse to press a GUI element 24 that has an appropriate label such as "List Sites." The  
25 Internet Research tool 102 detects the mouse event, contacts the web server 126, and retrieves a list of sites that match the selected topics. The list of sites is transmitted to the window 132 (step 44), and displayed as the scrolling list 25 in the right area of the user interface window below the third-level pull-down menu 22c (as shown in FIGS.  
4A-4C). The site list is implemented as an HTML file.

30 When a pre-selected web site 40 is chosen from the scrolling site list 25,

WO 99/49396

PCT/US99/06333

- 11 -

the web server 126 transmits a file that contains a brief description of the site, called a preview, to the local computer where the preview is displayed in target frame 26a in window 132 (step 46). The preview helps the user determine whether to access the 5 pre-selected web site 40 through a hypertext link appearing in window 132 as title 27. Thus, choosing pre-selected web site 40 and accessing pre-selected web site 40 are separated in time by previewing the site. The user has the choice of spending as much or as little time reading the preview as necessary and desired.

Internet research tool 102 utilizes a plugin. A plugin is a computer 10 software program that extends a browser's capabilities. Unlike a Java applet, a plugin allows processes to save data to a hard disk. Netscape and Microsoft provide plugin protocols and instructions to build plugins. Internet research tool 102 has one plugin written in C++ with customized computer code to extend the capabilities of Microsoft Corporation's Internet Explorer™ and Netscape Communication Corporation's 15 Navigator™ browsers. The plugin detects which browser is installed, and due to unique requirements of these browser products, only executes the code that is compatible for the installed browser. The plugin allows for drag and drop operations, gets bibliographic information, and saves the information to a local storage medium.

Internet research tool 102 maintains a data structure that contains the pre- 20 selected web site names and their respective URLs. When a user selects a pre-selected web site, Internet research tool 102 reads the data structure and gets the URL for the pre-selected web site selected by the user. The pre-selected web site is accessed using a plugin. The plugin is instantiated using JavaScript. A site name, site author, and URL are passed to the plugin via arguments in an EMBED call.

25 Internet research tool 102 takes advantage of the capability that multiple plugin instances share common memory. Global data (e.g., URL information) and a pointer to the main plugin are stored in common memory. Multiple instances of the plugin can come into and out of existence and share information, and temporary instances can initiate main plugin operations.

30 Internet research tool 102 instantiates the plugin on three different

WO 99/49396

PCT/US99/06333

- 12 -

occasions. First, selecting a web site while in search mode 12 causes a "site-selection" instance of the plugin to be embedded into an invisible frame. This generates a plugin shell, if one does not exist. The plugin shell maintains various global variables shared by other instantiations of the plugin. Other global variables produced by the site-selection instance maintain information about the last pre-selected web site visited. These variables are used by a "main" plugin instance and are attached to clippings as the clippings are generated.

A second occasion for instantiating the plugin is when switching to collect mode 14. This instantiation produces the "main" instance of the plugin, if one does not exist. A plugin shell is also produced at this time if one does not presently exist. The plugin generates objects that show interface elements, control disk access, handle drag operations, and maintain a list of clippings.

Internet research tool 102 detects when GUI element switch 20 changes state by using common user interface techniques, for example when a mouse click occurs while the mouse 110 is in a particular location. When GUI element switch 20 changes state to COLLECT, the Java applet Switcher.java 20' initiates the loading of the plugin. User interface collection tools, devoted to clipping and saving information, are displayed in the frame 26b in the right portion of window 132. The content in target frame 26a, displayed in the left portion of window 132, remains unchanged.

A third occasion for instantiating the plugin is when research tool 102 begins to produce a clipping. Both the plugin shell and "main" plugin objects should already exist. The "main" plugin object is retained as a global pointer in the plugin shell, so that its functions may be called.

In collect mode 14, the research tool acquires text (step 48), acquires images (step 50), and acquires audio data (step 52) from pre-selected web site 40. Text, images, and audio data can be collected into separate clipping files (18a, 18b, and 18c) and stored on disk 112. Text, images, and audio data can also be collected into the same clipping file.

WO 99/49396

PCT/US99/06333

- 13 -

To produce a text clipping in a PC/Netscape environment, the user highlights the desired text and clicks on the appropriate user-interface button. The collect area displays the clipping type, text clipping name, site author, site name, and 5 a notes box 37 where the user adds his or her own notes (as shown in FIG. 4B).

Research tool 102 automatically stores the clipping in an HTML clipping file 18 located locally on disk 112 and displays the human-perceptible form of the clipping 18' in target frame 26a. The user can edit the clipping file 18 (step 54), list 10 all clipping files 18 that exist in a file on disk 112 (step 56), and simply view the human-perceptible form of the clipping 18' (step 58). Along with the clipped data (e.g., text or image), the clipping name and type, the site name, author, and URL, and user notes are also displayed.

Other embodiments are within the scope of the following claims. Internet research tool 102 can be adapted to any field of interest, including business topics, 15 hobbies, and legal fields of interest. Internet research tool 102 may also clip video images from the Internet. User interface and implementation details differ depending on the browser installed on the computer system 100.

Having described preferred embodiments of the invention, it will now become apparent to one of skill in the art that other embodiments incorporating its 20 concept may be used. It is felt, therefore, that this invention should not be limited to the disclosed embodiment, but rather should be limited only by the spirit and scope of the appended claims.

WO 99/49396

PCT/US99/06333

- 14 -

What is claimed is:

1. A computer-implemented method for conducting research on the Internet, comprising:
  - accessing a web page on a computer network;
  - 5 providing a view of the web page on a computer monitor;
  - clipping a portion of data provided in the view;
  - saving the clipped portion of data to a file using the same graphical user interface that provides the view;
  - automatically adding a reference annotation corresponding to the web page
- 10 in the view to the file; and
- producing a human-perceptible document from the clipped data.
2. The method of claim 1, further comprising:
  - providing a switch to change from a search mode that accesses the web page to a collect mode that clips a portion of data provided in the view.
- 15 3. The method of claim 2, further comprising:
  - replacing a set of search tools with a set of collection tools in the same screen space.
4. The method of claim 3, wherein the search tools provide a hierarchical presentation of a plurality of topics and continuously display a selected hierarchical
- 20 path.
5. The method of claim 1, wherein the human-perceptible document is output to a frame in a web browser displayed on the computer monitor.
6. The method of claim 1, wherein the clipped portion of data includes more

WO 99/49396

PCT/US99/06333

- 15 -

than one feature of the web page.

7. The method of claim 6, wherein the features include multi-media data.
8. The method of claim 1, further comprising:  
adding a user-generated annotation to the file.
- 5 9. The method of claim 1, wherein the reference annotation includes a  
document name, an author, a URL, or a date.
10. The method of claim 1, wherein the reference annotation is bibliographic  
data.
11. The method of claim 1, further comprising:  
10 allowing access only to pre-selected web pages; and  
displaying a preview of the web page prior to accessing the web page.
12. A computer-implemented method for conducting research on the Internet,  
comprising:  
choosing a topic from a list of topics;  
15 locating data stored on a first computer medium attached to a remote  
computer, wherein the data is associated with the topic;  
displaying the data on a local computer monitor;  
switching from a search mode to a collection mode;  
selecting a portion of the data;  
20 saving the portion of the data in a file on a second computer medium; and  
saving a site reference with the data in the file on the second computer  
medium.

WO 99/49396

PCT/US99/06333

- 16 -

13. The method of claim 12, further comprising:  
editing the file to include user-supplied data.
14. The method of claim 12, further comprising:  
remaining within the same user interface boundary while locating,  
selecting, and saving data.
15. A computer program, residing on a computer-readable medium,  
comprising instructions causing a computer to:  
present a selection of web pages on a computer monitor;  
access one web page in the selection through a computer network;  
10 display the one web page in a frame on the computer monitor;  
clip a portion of data from the frame;  
save the clipped portion of data to a file;  
add an annotation to the file; and  
produce a human-perceptible document from the clipped data.
- 15 16. The computer program of claim 15, further comprising instructions  
causing a computer to:  
annotate the clipped data with a bibliographic reference.
17. The computer program of claim 15, further comprising instructions  
causing a computer to:  
20 display a description of the web page prior to accessing the web page.
18. A computer program, residing on a computer-readable medium,  
comprising instructions causing a computer to:  
choose a topic from a hierarchical list of topics while displaying a  
hierarchical path;

WO 99/49396

PCT/US99/06333

- 17 -

locate data stored on a first computer medium attached to a computer,  
wherein the data is associated with the topic;  
display the data on a computer monitor;  
switch from a search mode to a collection mode;  
5 select a portion of the data;  
save the portion of the data to a file on a second computer medium; and  
save a site reference with the data in the file on the second computer  
medium.

19. The computer program of claim 18, further comprising instructions  
10 causing a computer to:  
display a preview of the web page prior to accessing the web page; and  
remain within one user-interface environment during the search mode and  
collect mode.
20. The method of claim 19, further comprising instructions causing a  
15 computer to:  
add user-supplied data to the file.

WO 99/49396

PCT/US99/06333

1/10

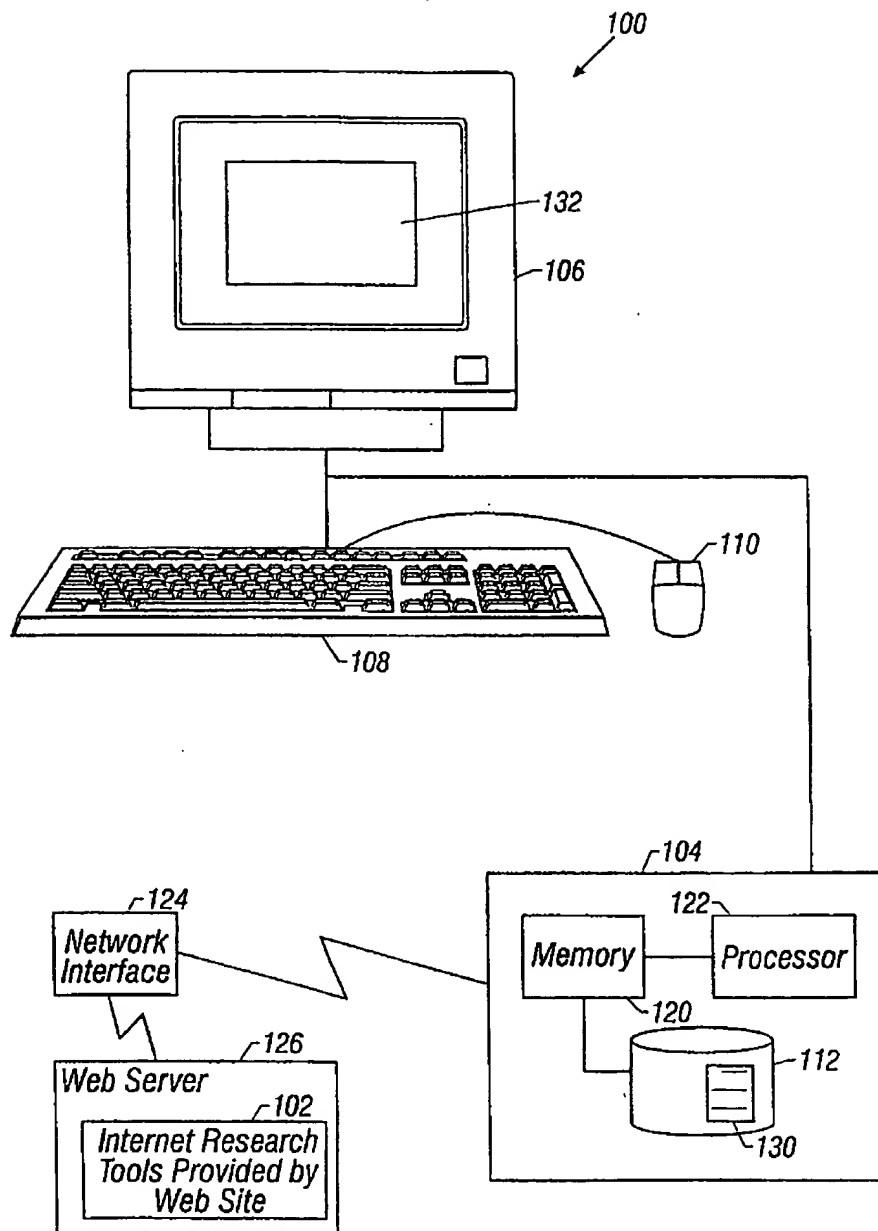


FIG. 1

SUBSTITUTE SHEET (RULE 26)

Printed from Mimosa 07/02/2003 12:41:03 Seite -20-

WO 99/49396

PCT/US99/06333

2/10

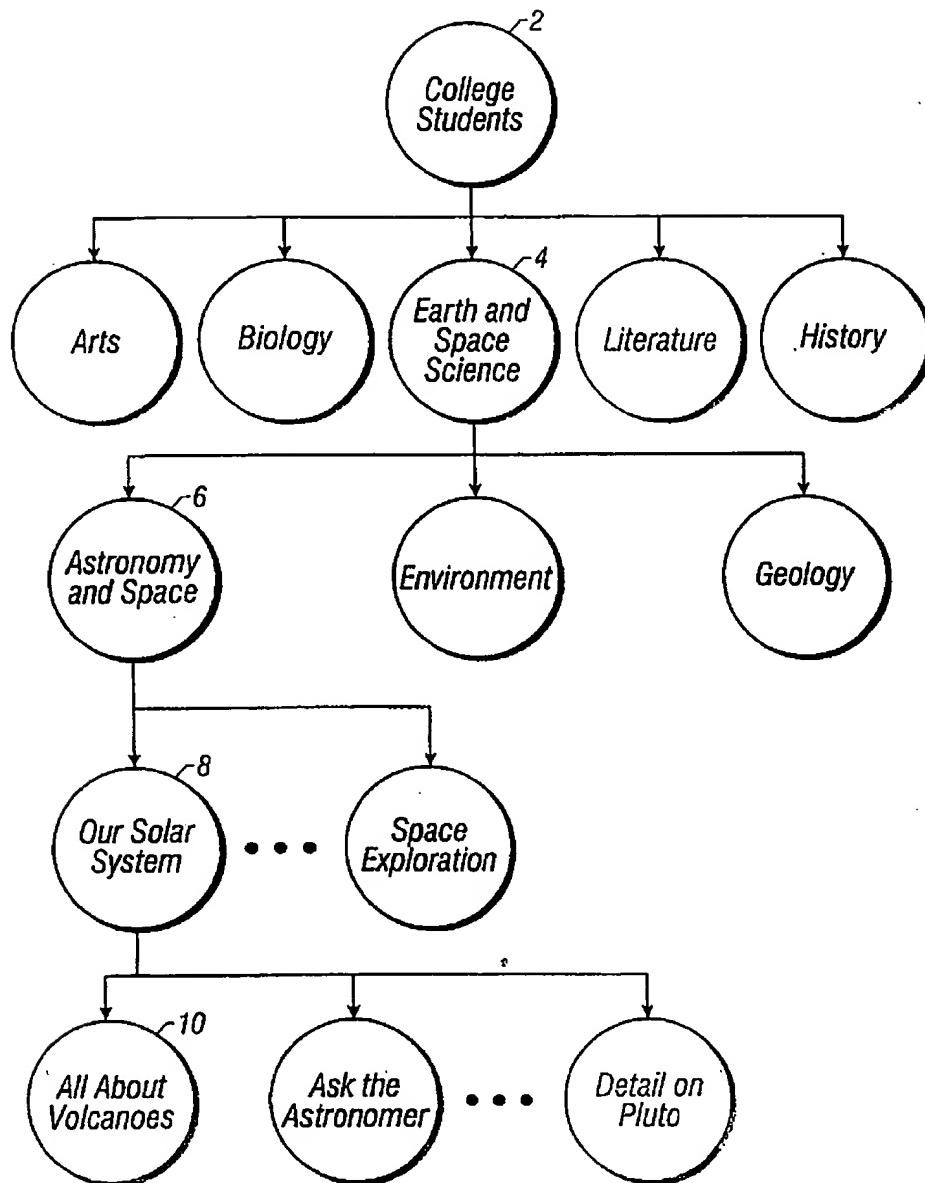


FIG. 2

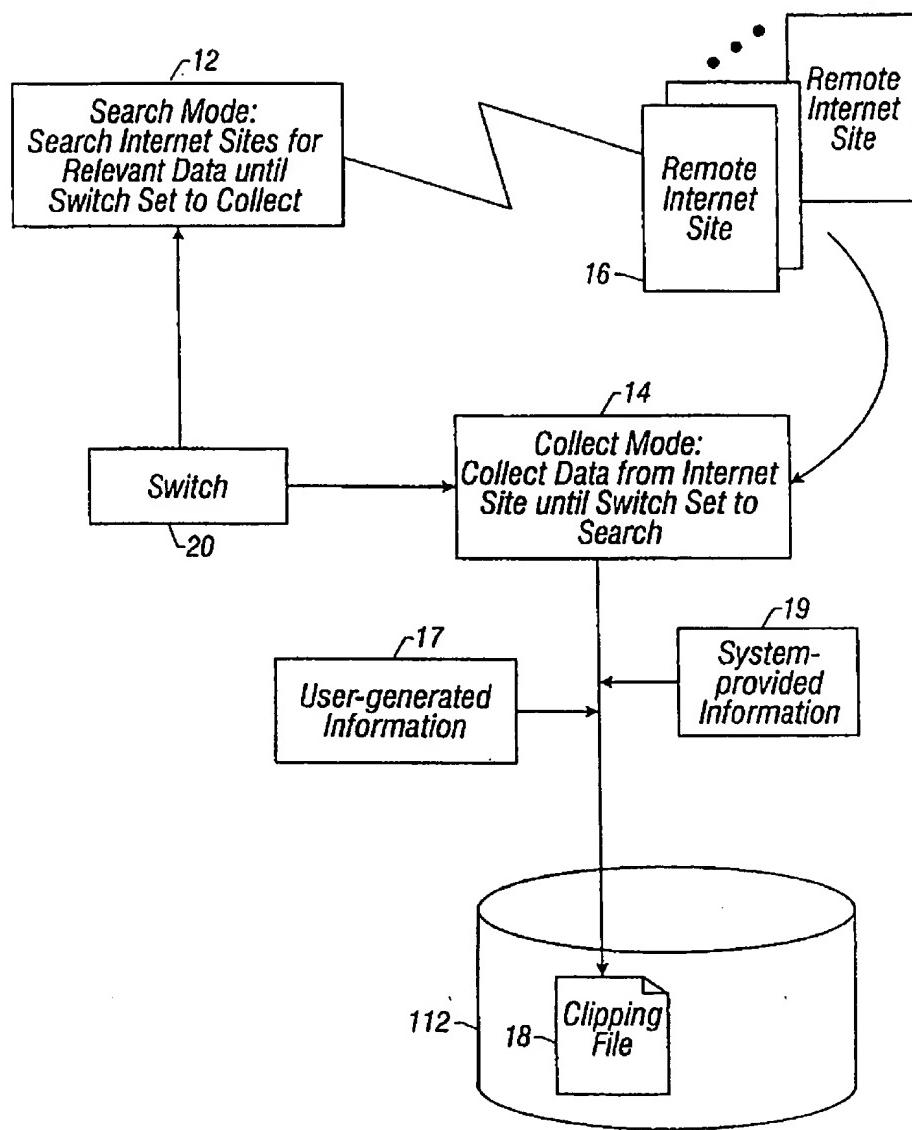
## SUBSTITUTE SHEET (RULE 26)

Printed from Mimosa 07/02/2003 12:41:03 Seite -21-

WO 99/49396

PCT/US99/06333

3/10



**FIG. 3**  
SUBSTITUTE SHEET (RULE 26)

WO 99/49396

PCT/US99/06333

4/10

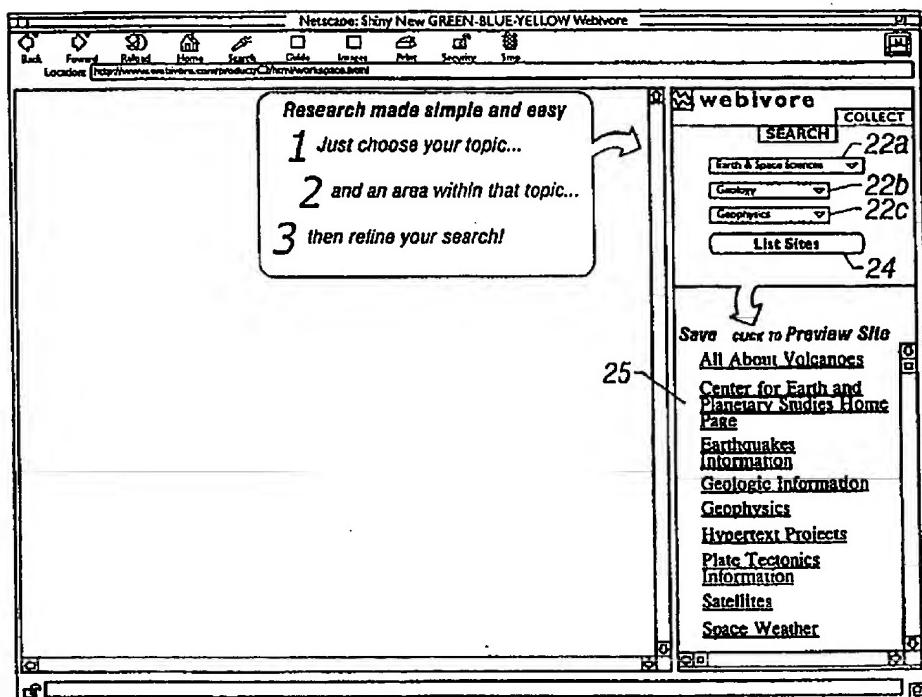


FIG. 4A

## SUBSTITUTE SHEET (RULE 26)

Printed from Mimosa 07/02/2003 12:41:03 Seite -23-

WO 99/49396

PCT/US99/06333

5/10

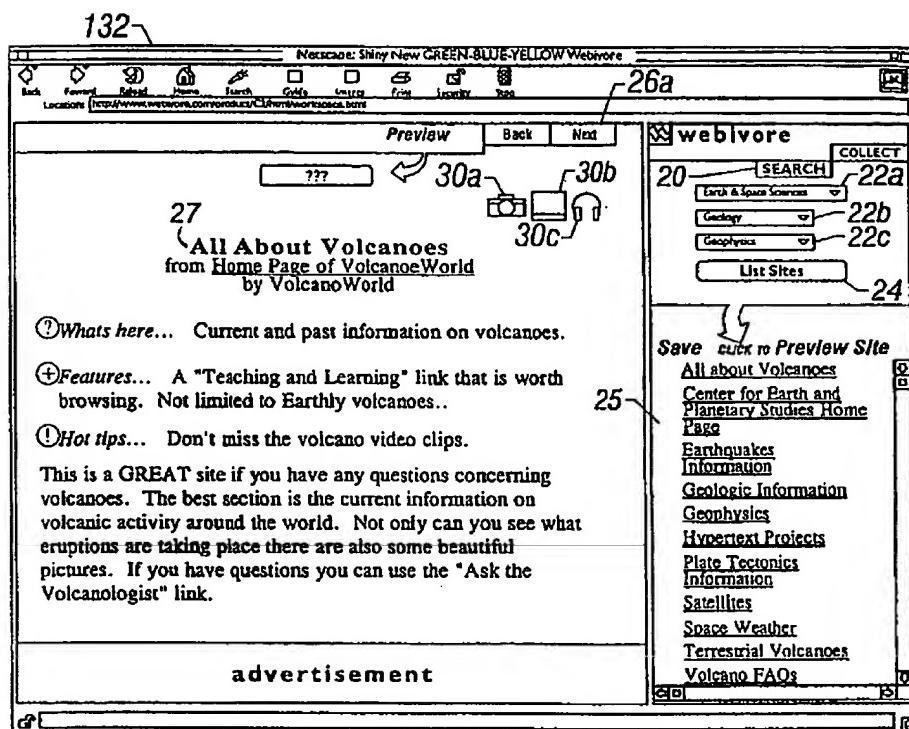


FIG. 4B

## SUBSTITUTE SHEET (RULE 26)

Printed from Mimosa 07/02/2003 12:41:03 Seite -24-

WO 99/49396

PCT/US99/06333

6/10

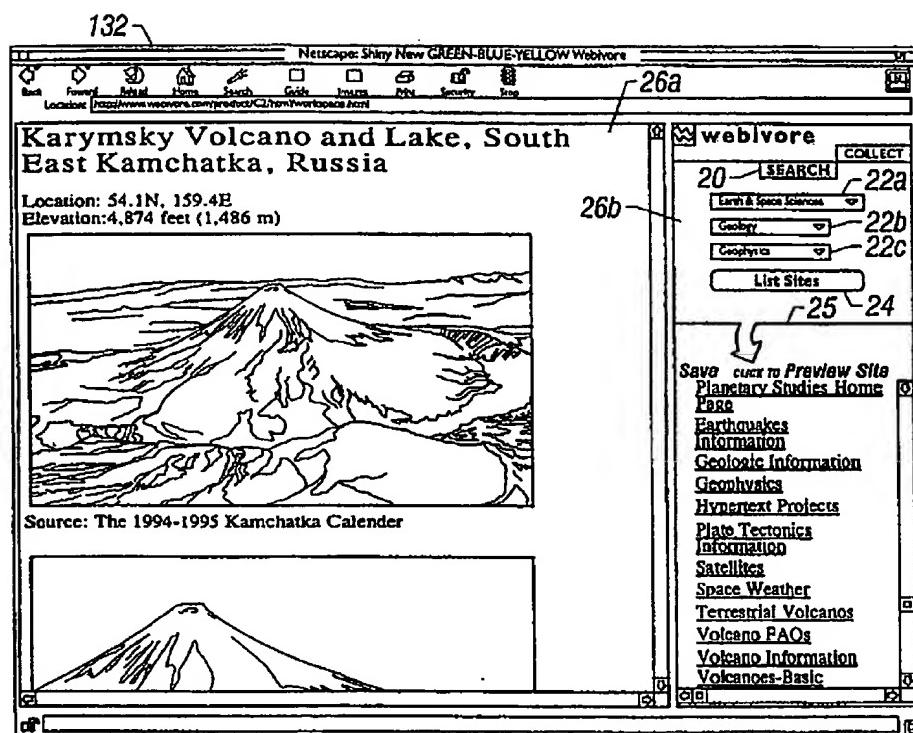


FIG. 4C

## SUBSTITUTE SHEET (RULE 26)

Printed from Mimosa 07/02/2003 12:41:03 Seite -25-

WO 99/49396

PCT/US99/06333

7/10

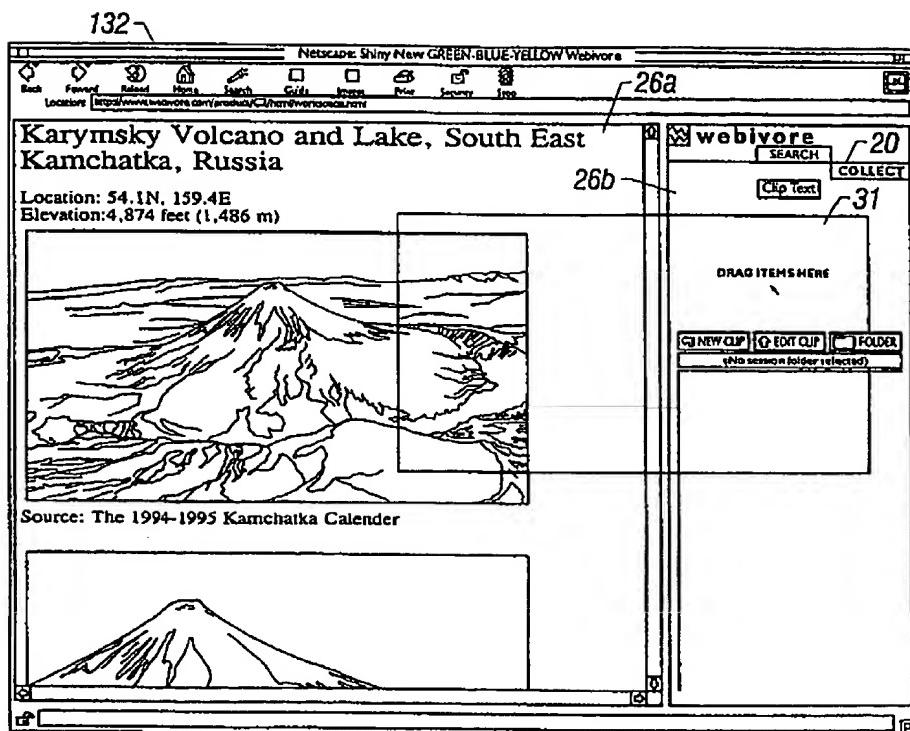


FIG. 4D

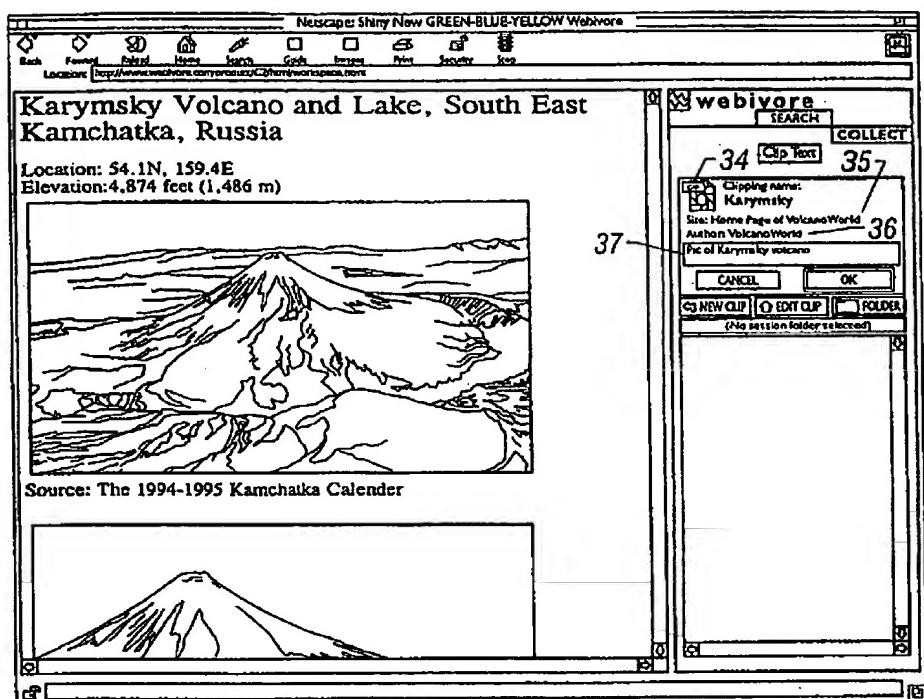
## SUBSTITUTE SHEET (RULE 26)

Printed from Mimosa 07/02/2003 12:41:03 Seite -26-

WO 99/49396

PCT/US99/06333

8/10



**FIG. 4E**

**SUBSTITUTE SHEET (RULE 26)**

Printed from Mimosa 07/02/2003 12:41:03 Seite -27-

WO 99/49396

PCT/US99/06333

9/10

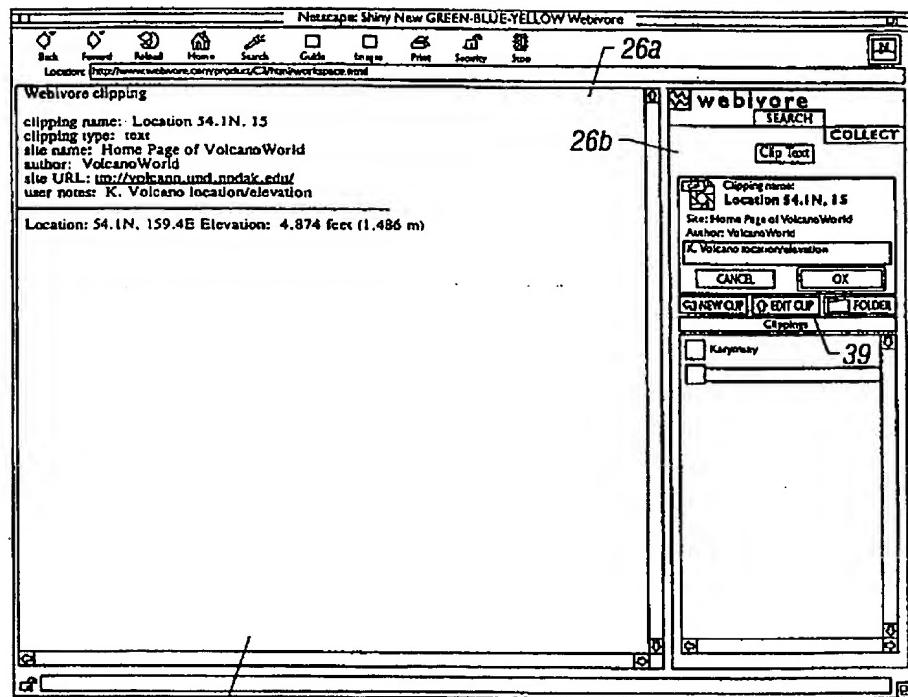


FIG. 4F

**SUBSTITUTE SHEET (RULE 26)**

Printed from Mimosa 07/02/2003 12:41:03 Seite -28-

WO 99/49396

PCT/US99/06333

10/10

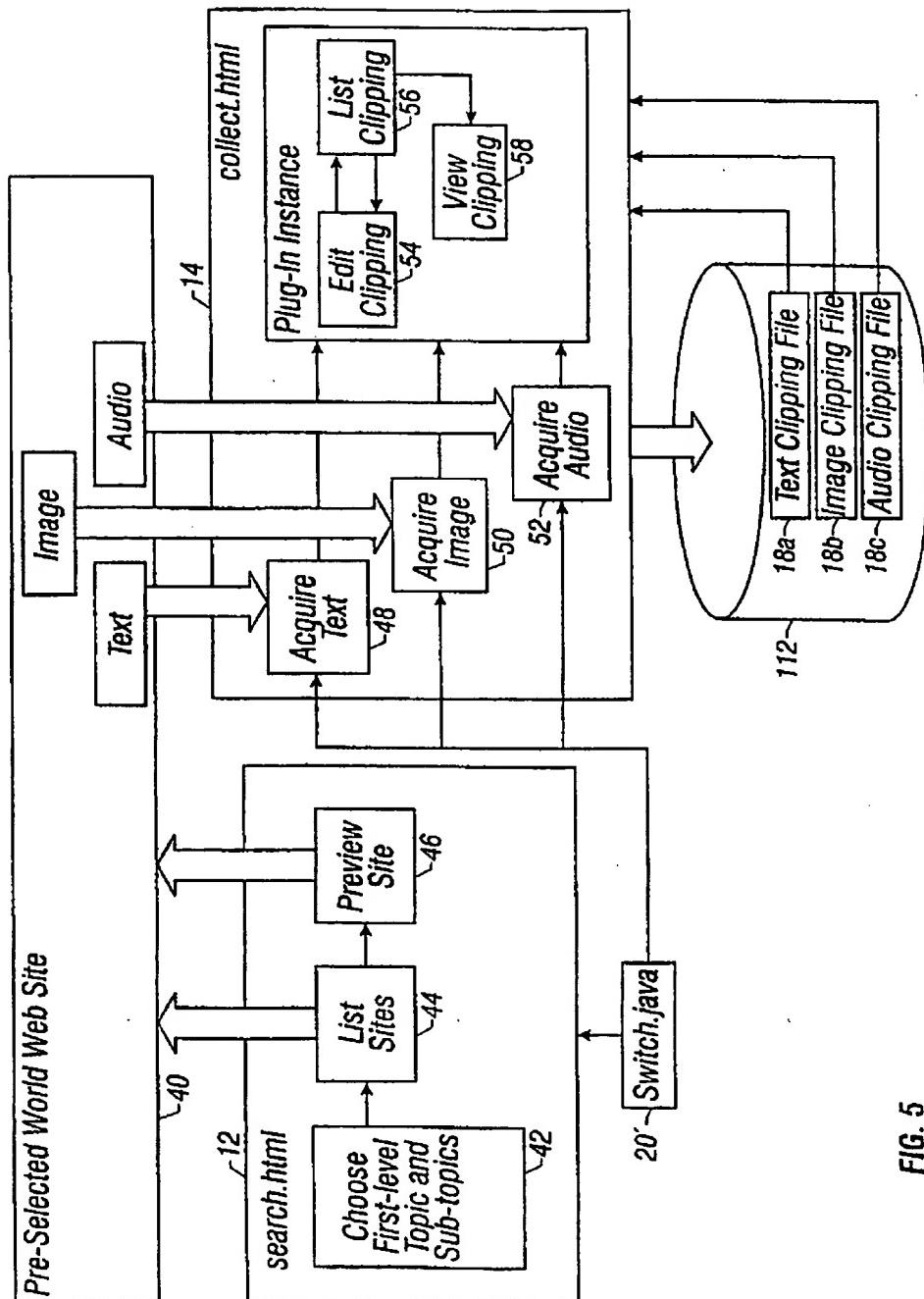
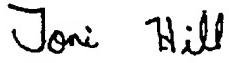


FIG. 5

SUBSTITUTE SHEET (RULE 26)

Printed from Mimosa 07/02/2003 12:41:03 Seite -29-

INTERNATIONAL SEARCH REPORT		International application No. PCT/US99/06333
<b>A. CLASSIFICATION OF SUBJECT MATTER</b> IPC(6) :G06F 13/00, 15/00; 17/22; 17/30; G06T 11/00 US CL :707/10, 501, 513, 515; 395/200.66 According to International Patent Classification (IPC) or to both national classification and IPC		
<b>B. FIELDS SEARCHED</b> Minimum documentation searched (classification system followed by classification symbols) U.S. : 707/10, 501, 513, 515; 395/200.66		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched NONE		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) APS		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 5,649,186 A (FERGUSON) 15 July 1997, see abstract, col. 1, lines 50 to col. 2, lines 21, fig. 1, 2, 4.	1, 12, 15, 18
Y	US 5,659,676 A (REDPATH) 19 August 1997, see abstract, col. 2, lines 20 to col. 3, lines 40	1, 12, 15, and 18
Y,P	US 5,822,539 A (VAN HOFF) 13 October 1998, see abstract, col. 2, lines 65 to col. 3, lines 40	1-20
Y,P	US 5,845,299 A (ARORA et al.) 01 December 1998, see abstract, col. 2, lines 20-61.	1-20
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "B" earlier document published on or after the international filing date "C" document which may throw doubt on priority claim(s) or which is cited to establish the publication date of another citation or other special reasons (as specified) "D" document referring to an oral disclosure, use, exhibition or other means "E" document published prior to the international filing date but later than the priority date claimed		
*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "A" document member of the same patent family		
Date of the actual completion of the international search	Date of mailing of the international search report	
29 APRIL 1999	15 JUN 1999	
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer THUY PARDO Telephone No. (703) 305-1091 	

Form PCT/ISA/210 (second sheet)(July 1992)\*

Printed from Mimosa 07/02/2003 12:41:03 Seite -30-

INTERNATIONAL SEARCH REPORT		International application No. PCT/US99/06333
C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y,P	US 5,870,770 A (WOLFE) 09 February 1999, see abstract, col. 3, lines 40 to col. 4, lines 29.	1-20

Form PCT/ISA/210 (continuation of second sheet)(July 1992)\*